Efficient Thresholding Technique Using Neural Networks (NN)

Mohammed Jahirul Islam

November 2008

Presentation Outline

- Image Thresholding

- Artificial Neural Network (ANN)

- NN-based Thresholding technique

- Training Data Preparation, Testing NN

- Observations and Criticisms

- Proposed Technique Flowchart

- Simulation Results

- Conclusions and Future Work

Image Thresholding

- Digital color image is represented by 24-bit (16- millions levels) or gray scale image (scanned document) by 8-bit (256 levels)

- The analysis of an image with that many levels might require complicated techniques and higher computational cost

- Reduce the image to a more manageable number of grey levels, usually two levels (binary image), and at the same time retain all necessary features of the original image

- **Thresholding is a way of solving this issue**

Image Thresholding

- Two broad categories
  - Global thresholding- Picks one value for the entire image
  - Local thresholding- different value for different pixels, adapative

- Selection of appropriate thresholding technique is application dependent.

- **Document Analysis is one of the important applications**
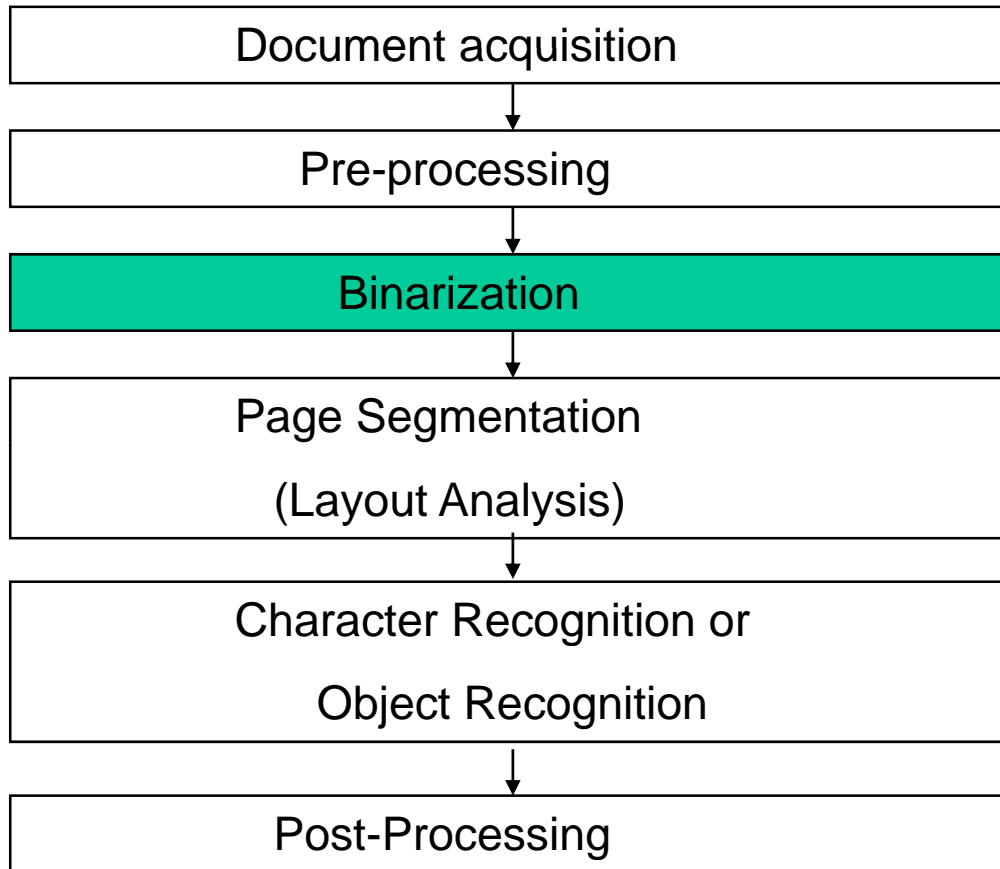
Document Analysis

- Speech is a sign system that is more natural than writing to humans

- Writing is considered to have made possible much of culture and civilization.

- Printed documents, such as newspapers, magazines and books, and in handwritten matter, such as found in notebooks and personal letters.

- **Document Analysis System converts a paper-based document into computerized form**

Document Analysis System (DAS)

- Recognize characters of a text block and identify non-text regions such as charts and images

- Advantages includes efficient document updates and revisions

- Most of the successes have come in constrained domains such as postal addresses, bank checks, and census forms.

Principles Stages of DAS

```
┌─────────────────────────────────┐
│      Document acquisition        │
└─────────────────────────────────┘
                 │
                 ▼
┌─────────────────────────────────┐
│         Pre-processing           │
└─────────────────────────────────┘
                 │
                 ▼
┌─────────────────────────────────┐
│          Binarization            │
└─────────────────────────────────┘
                 │
                 ▼
┌─────────────────────────────────┐
│        Page Segmentation         │
│        (Layout Analysis)         │
└─────────────────────────────────┘
                 │
                 ▼
┌─────────────────────────────────┐
│     Character Recognition or     │
│        Object Recognition        │
└─────────────────────────────────┘
                 │
                 ▼
┌─────────────────────────────────┐
│         Post-Processing          │
└─────────────────────────────────┘
```

Limitation of DAS

- Giant steps have been made in the last decade, both in terms of technological supports and in software products to provide computerized DAS.

- Character recognition (OCR) contributes to this progress by providing techniques to convert large volumes of data automatically.

- There are so many papers and patents advertising recognition rates as high as 99.99%; this gives the impression that automation problems seem to have been solved.

- **What if the document is composite and degraded?**

Challenges in DAS

- Performance problems subsist on composite paper documents with non-uniform background.

- Non-uniform background is caused by watermarks and complex patterns used in printing security documents

- Success of converting documents with complex backgrounds depends on
  - Eliminating background by thresholding
  - Correctness of page segmentation

- **Main challenge is Image Binarization**

Literature Review

- Performance of thresholding depends on the type of document, image illumination, contrast and complexity of the background

- Trier and Jain [1] compared several local and global thresholding technique and their respective character recognition rate.
  - Niblack [2] local adaptive method produced the best
- Sahoo et al. [3] compared 20 global thresholding methods
  - Otsu [4] outperformed all other methods

- All thresholding techniques do not perform well on all images
  - Most make some assumptions about the images to be used which limit their performance to such images

Literature Review

- Yasser [5] developed NN-based technique for thresholding composite digitized documents with complex background

  - Passports, bank cheques, ID cards and images from magazines and scanned synthetic images printed on complex background

- **What is Artificial Neural Network?**

Artificial Neural Network (ANN)

- Powerful data modeling tool, represents complex input/ output relationships

- Resembles human brain in acquiring knowledge through learning and storing knowledge within inter-neuron connection strength, Weights.

How does ANN work?

- ANNs area adjusted or trained so that a particular input leads to a specific desired or target output
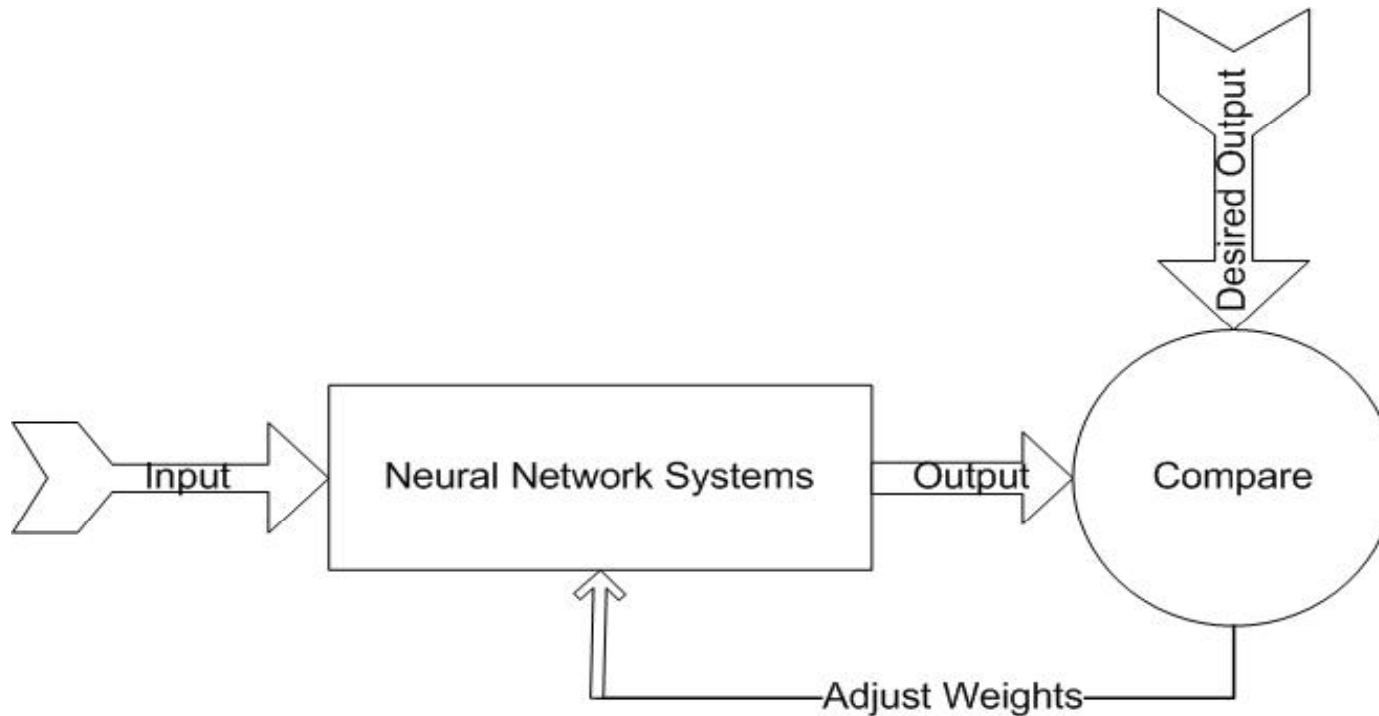
Multi-Layer Perceptron (MLP)

- Most common NN model
- Uses supervised training methods to train the NN
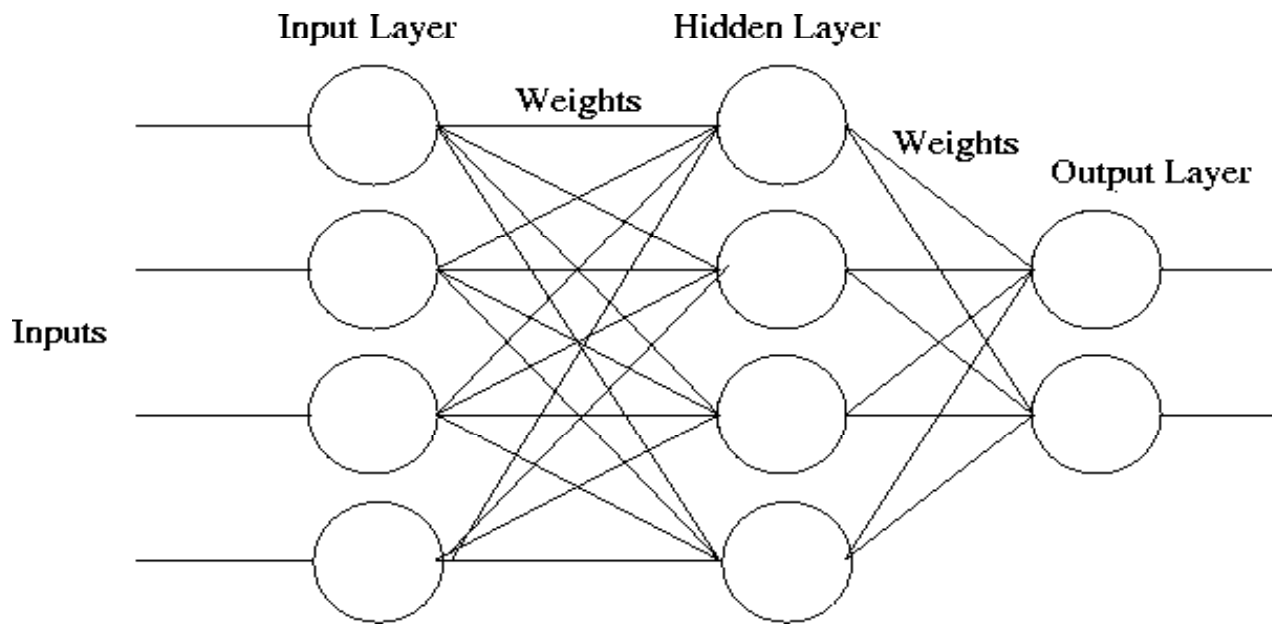
NN and Thresholding

- Very few researchers have investigated the use of NNs in image thresholding.

- Koker and Sari [8] use NNs to automatically select a global threshold value for an industrial vision system

- Papamarkos [9] produced a local thresholding method using the Kohonen SOM classifier to define the two bi-level classes in order to reduce the character blurring effect in blurred documents

NN-based Thresholding

- Poor contrast, non-uniform illumination, complex background patterns and non –uniformly distributed background is a challenging problem in thresholding of document images

- NN-based algorithm uses statistical and textural feature measures to obtain a feature vector from a pixel window of size (2n+1) x (2n+1), where n>=1

- Uses MLP NN to train the network and adjust the weights and then classify each pixel in the image

Statistical Texture Measures

- Statistical textural measures are useful in characterizing the set of neighborhood values of pixels.

- Features:
  - Pixel value
  - Mean
  - Standard Deviation
  - Smoothness
  - Entropy
  - Skewness
  - Kurtosis
  - Uniformity

Training Data Preparation

- Load an image
- Select a pixel in the image
- Click on object or background button for the selected pixel
- All the 8 features are calculated
- Save it in a file as a feature vector
- Repeat the process for random points and another image

Training Data Preparation

Training the Network

- Input Layer- No. of features in a feature vector, for example 8
- Hidden layer- (Input+output)*2/3
- Output layer- 1 (Object 0, background 1)
- Weights=(Input*Hidden)+Hidden units

- Use supervised training methods to train the NN
- Training sequence involves forward phase and backward phase

- Forward phase estimates the error and backward phase modifies the weights to decrease the error

Testing the NN

- Weights are used in classification phase
- Image data feature vectors are extracted from each pixel and its neighborhoods, fed into the network that performs classification and assign a number 0 or 1

Observations and Criticisms

- Feature vector (all the 8 features) inputs to the NN

- More features used slower the feature extraction process

- Window size affects the speed, larger the window size slower the feature extraction process

- Window size 5x5 used in this case

- **Should we use all 8 features?**
- **Is it possible to have a combination with minimum features and higher recognition rate? If so, What combination?**

Objectives of the Research

- Reduce the number of features and search the combination that is minimum but provide the same or better recognition rate

- Validate the combination by testing on more images

- Propose an efficient thresholding technique using the combination of minimum features

Features Combination

- How many different combinations possible using 8 features?

$$^8C_8 = 1 \qquad\qquad ^8C_4 = 70$$

$$^8C_7 = 8 \qquad\qquad ^8C_3 = 56 \qquad ^8C_7 = 8 = \begin{cases} 1,2,3,4,5,6,7 \\ 1,2,3,4,5,6,8 \\ 1,2,3,4,5,7,8 \\ 1,2,3,4,6,7,8 \\ 1,2,3,5,6,7,8 \\ 1,2,4,5,6,7,8 \\ 1,3,4,5,6,7,8 \\ 2,3,4,5,6,7,8 \end{cases}$$

$$^8C_6 = 28 \qquad\qquad ^8C_2 = 28$$

$$^8C_5 = 56 \qquad\qquad ^8C_1 = 8$$

- **Total: 255 combination without repetition**

Flowchart of the Proposed Process

```
Document          NN-based          Commercial        Recognized
Image      →      Thresholding  →   OCR (ABBYY)   →   Characters
                                                       (RC)
                      │                                   │
                      ▼                                   ▼
                  Weights- 255      Expected          Error Rate
                  Combinations      Characters   →    ={(EC-RC)/
                                    (EC)              EC}*100
```
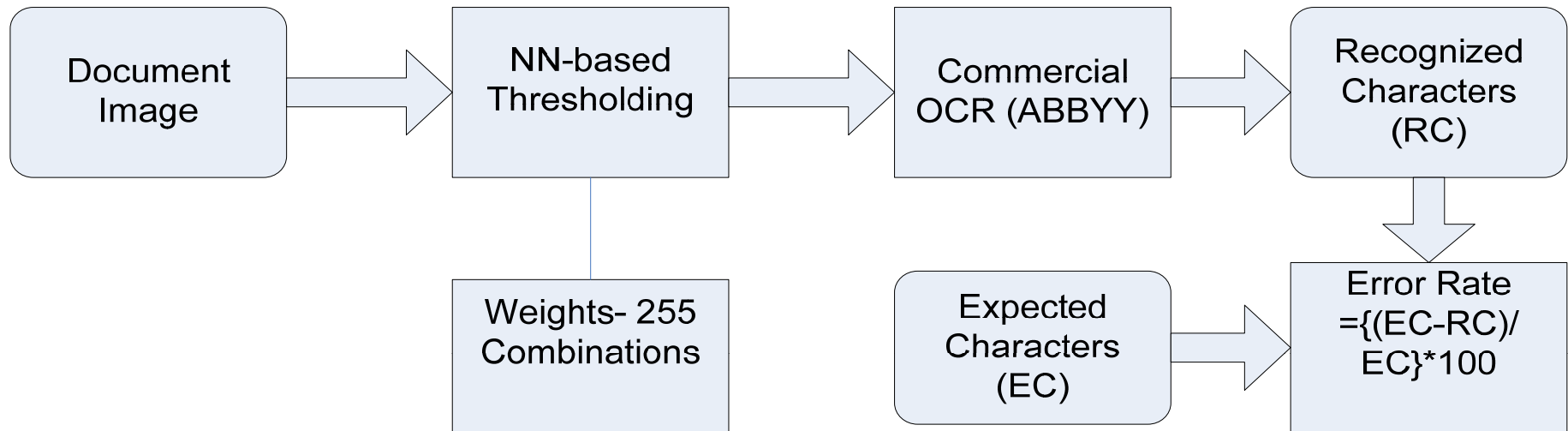
Observations of the Proposed Process

- Different feature vector have different weights
  - 255 features combination, 255 Weight vector

- For each document image 255 OCR output, 255 error rate

- Compare the error rate and picks up the best combination

- Repeat the same process for simple, moderate and complex background document images

- **Select the minimum feature combination with high recognition rate**

Simulation Results- Sample Testing Images

Health

Health is defined as a state of complete physical, social and mental well-being, and not merely the absence of disease or infirmity. Within the context of health promotion, health has been considered less as an abstract state and more as a means to an end which can be expressed in functional terms as a resource which permits people to lead an individually, socially and economically productive life. Health is a resource for everyday life, not the object of living. It is a positive concept emphasizing social and personal resources as well as physical capabilities.

The Underground Railroad

In the days before and during the American Civil War, Ontario served as the final stop on the underground railroad, a network of secret routes and safe houses that allowed enslaved African-Americans to escape to freedom in Canada.

Walk in the footsteps of history along the African Canadian Heritage Route from Windsor, where you can visit John Freeman Walls' 1846 log cabin, that served as a terminal on the Underground Railroad. For a further window into the past, walk among the artifacts and images at the Amherstburg's North American Black Historical Museum. Stroll the streets of North Buxton, Canada's first Black settlement – home to many historic buildings and a museum that recounts the area's proud story of early growth and self-sufficiency.

Rail Road

Arnold
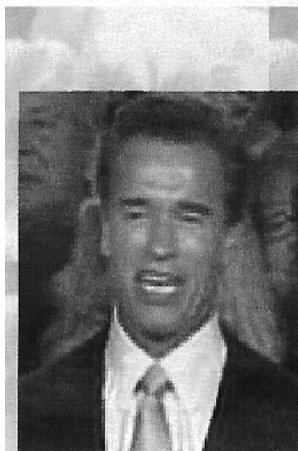
In early 2003, Californians pointed fingers as their state struggled with a $38 billion budget deficit and a continuing energy crisis. Republicans set their sights on Democratic Gov. Gray Davis, attempting to make him the second governor in U.S. history to be recalled. On October 7, the majority of voters decided to oust Davis, then chose a successor from among 135 candidates. One of Hollywoods own took Davis place bodybuilder-turned-actor Republican Arnold Schwarzenegger.

Comparative Statement

Image	Total Chars	Features and Recognition Rate (%)			Commercial OCR
		1, 2, 5	**1, 2, 6**	**1, 5, 6**	**ABBYY (7.0)**
Health	476	**99.79 (1)**	99.79 (1)	99.58 (2)	99.37 (3)
Rail Road	654	**99.85 (1)**	99.54 (3)	98.32 (11)	99.24 (5)
Arnold	405	**99.51 (2)**	99.01 (4)	96.30 (15)	96.79 (13)

Features:
1. Pixel   5. Entropy
2. Mean   6. Skewness
3. Std. Dev.   7. Kurtosis
4. Smoothness   8. Uniformity

Comparative Statement

Image	Total Chars	Features and Recognition Rate (%)	
		1, 2, 5	**1, 2, 6**
Niagra Falls	339	**100 (0)**	99.41 (2)
Chretien	426	**99.77 (1)**	99.53 (2)
George	414	**99.52 (2)**	99.28 (3)
Volcanos	348	**100 (0)**	99.71 (1)
Cats	585	**99.66 (2)**	99.66 (2)

Segmented Image- 1, 2, 5
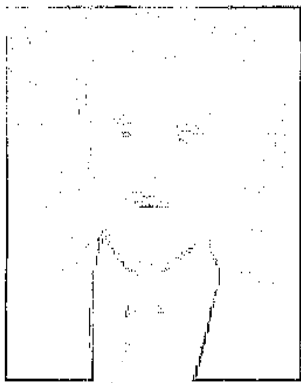
In early 2003, Californians pointed fingers as their state struggled with a $38 billion budget deficit and a continuing energy crisis. Republicans set their sights on Democratic Gov. Gray Davis, attempting to make him the second governor in U.S. history to be recalled. On October 7, the majority of voters decided to oust Davis, then chose a successor from among 135 candidates. One of Hollywoods own took Davis place bodybuilder-turned-actor Republican Arnold Schwarzenegger.

Health is defined as a state of complete physical, social and mental well-being, and not merely the absence of disease or infirmity. Within the context of health promotion, health has been considered less as an abstract state and more as a means to an end which can be expressed in functional terms as a resource which permits people to lead an individually, socially and economically productive life. Health is a resource for everyday life, not the object of living. It is a positive concept emphasizing social and personal resources as well as physical capabilities.

The Underground Railroad

In the days before and during the American Civil War, Ontario served as the final stop on the underground railroad, a network of secret routes and safe houses that allowed enslaved African-Americans to escape to freedom in Canada.

Walk in the footsteps of history along the African Canadian Heritage Route from Windsor, where you can visit John Freeman Walls' 1846 log cabin, that served as a terminal on the Underground Railroad. For a further window into the past, walk among the artifacts and images at the Amherstburg's North American Black Historical Museum. Stroll the streets of North Buxton, Canada's first Black settlement – home to many historic buildings and a museum that recounts the area's proud story of early growth and self-sufficiency.

OCR Output- Expected Vs. 1,2,5

- in early 2003, Californians pointed fingers as their state struggled with a $38 billion budget deficit and a continuing energy crisis. Republicans set their sights on Democratic Gov. Gray Davis, attempting to make him the second governor in U.S. history to be recalled. On October 7, the majority of voters decided to oust Davis, then chose a successor from among 135 candidates. One of Hollywoods own took Davis place bodybuilder-turned-actor Republican Arnold Schwarzenegger.
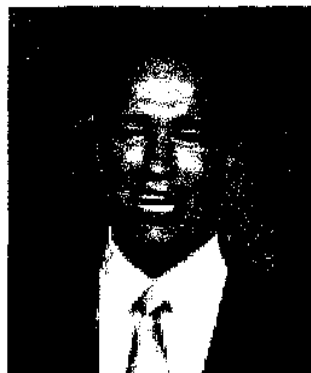
- In early 2003, Californians pointed fingers as their state struggled with a $38 billion budget deficit and a continuing energy crisis. Republicans set their sights on Democratic Gov. Gray Davis, attempting to make him the second governor in U.S. history to be recalled. On October 7, the majority of voters decided to oust Davis, then chose a successor from among 135 candidates. One of Hollywoods own took Davis place bodybuilder-turned-actor Republican Arnold Schwarzenegger

1, 2, 5

Expected

OCR Output- Expected Vs. ABBYY

- in early 2003, Caii form ans pointed fingers as their state struggled with a $38 billion budget deficit and a continuing energy crisis. Republicans set their sights on Democratic Gov. Gray Davis, attempting to make him the second governor in U.S. history to be recalled. On October 7, the majority of voters decided to oust Davis, then chose a successor from among 135 candidates. One of^H^llywoods own took Davi solace bodybupder-turned-aci Republican Arnold Schwarzenegger.

- In early 2003, Californians pointed fingers as their state struggled with a $38 billion budget deficit and a continuing energy crisis. 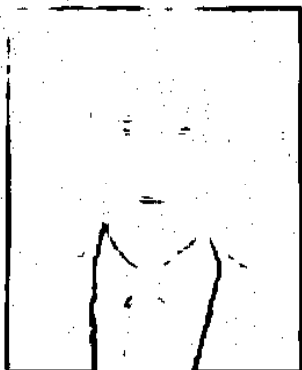Republicans set their sights on Democratic Gov. Gray Davis, attempting to make him the second governor in U.S. history to be recalled. On October 7, the majority of voters decided to oust Davis, then chose a successor from among 135 candidates. One of Hollywoods own took Davis place bodybuilder-turned-actor Republican Arnold Schwarzenegger

ABBYY   **Expected**

Otsu Output

In early 2003,
Californians
pointed fingers
as their state
struggled with a
$38 billion
budget deficit
and a continuing
energy crisis.
Republicans set
their sights on
Democratic Gov. Gray Davis, attempting to
make him the second governor in U.S.
history to be recalled. On October 7, the
majority of voters decided to oust Davis,
then chose a successor from among 135
candidates. One of Hollywoods own took
Davis place bodybuilder-turned-actor
Republican Arnold Schwarzenegger.

Segmented

- in early 2003, Cain form'ans granted fingers, a?tfite''r state st7i|g1ed with a' $38 billion budget deficit and a continuing energy crisis. RWujpicans set th^w sights on Democratic Gov. Gray Davis, attempting to make him the second governor in U.S. history to be recalled. On October 7, the majority of voters decided to oust Davis, then chose a succJSs€r f rom amonjL 1??
  ft   one, otJhB 1 ywoods .oiOTit&jk bodyl&«:der-turned-£rfr

OCR Output

Niblack Output

In early 2003, Californians pointed fingers as their state struggled with a $38 billion budget deficit and a continuing energy crisis. Republicans set their sights on Democratic Gov. Gray Davis, attempting to make him the second governor in U.S. history to be recalled. On October 7, the majority of voters decided to oust Davis, then chose a successor from among 135 candidates. One of Hollywoods own took Davis place bodybuilder-turned-actor Republican Arnold Schwarzenegger.

- in early 2003, **Califbrnians** pointed fingers as their state struggled with a $38 billion budget deficit and a continuing energy crisis. Republicans set their sights on Democratic Gov. **Cray** Davis, attempting to **aake Ma** the second governor In U.S. history to be recalled, on October 7, the **aajority** of voters decided to oust Davis, then chose a successor **froa aaong** 135 candidates, one of Hollywood* own **tnok** Davis place bodybuilder-turned-actor Republican Arnold Schwarzenegger.

Segmented

OCR Output

Simulation Results- 1,2,5

Health is defined as a state of complete
physical, social and mental well-being,
and not merely the absence of disease
or infirmity. Within the context of health
promotion, health has been considered
less as an abstract state and more as a
means to an end which can be
expressed in functional terms as a
resource which permits people to lead
an individually, socially and
economically productive life. Health is a
resource for everyday life, not the
object of living, it is a positive concept
emphasizing social and personal
resources as well as physical
capabilities.

Segmented

OCR Output

Simulation Results- 1,2,5

Segmented

The Underground Railroad

In the days before and during the American Civil War, Ontario served as the final stop on the underground railroad, a network of secret routes and safe houses that allowed enslaved African-Americans to escape to freedom in Canada.

Walk in the footsteps of history along the African Canadian Heritage Route from Windsor, where you can visit John Freeman Walls' 1846 log cabin, that served as a terminal on the Underground Railroad. For a further window into the past, walk among the artifacts and images at the Amherstburg's North American Black Historical Museum. Stroll the streets of North Buxton, Canada's first Black settlement – home to many historic buildings and a museum that recounts the area's proud story of early growth and self-sufficiency.

The Underground Railroad

in the days before and during the American civil War, Ontario served as the final stop on the underground railroad, a network of secret routes and safe houses that allowed enslaved African-Americans to escape to freedom in Canada.

Walk in the footsteps of history along the African Canadian Heritage Route from Windsor, where you can visit John Freeman Walls' 1846 log cabin, that served as a terminal on the Underground Railroad. For a further window into'the past, walk among the artifacts and images at the Amherstburg's North American Black Historical Museum, stroll the streets of North Buxton, Canada's first Black settlement - home to many historic buildings and a museum that recounts the area's proud story of early growth and self-sufficiency.

OCR

Simulation Results- 1,2,5

Document Image
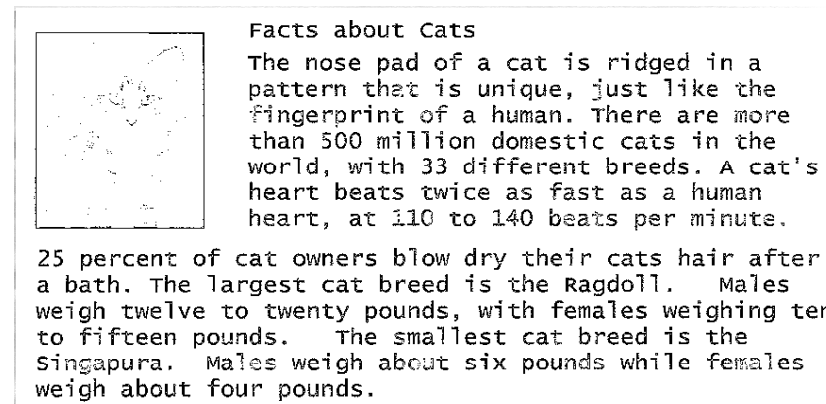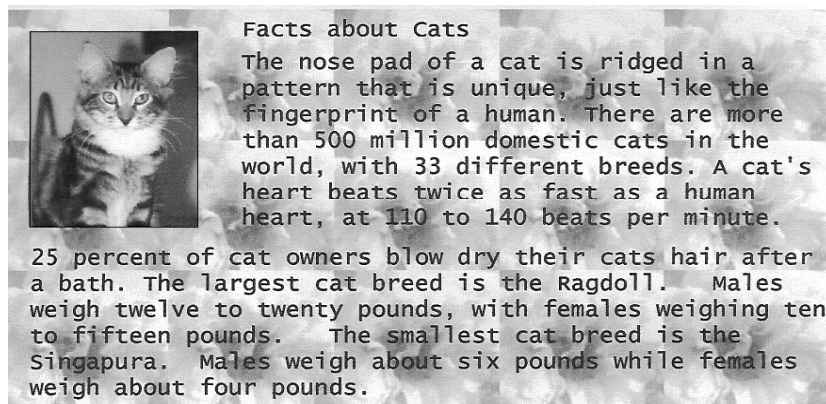
Segmented

OCR

Facts about Cats

The nose pad of a cat is ridged in a pattern that is unique, just like the fingerprint of a human. There are more than 500 million domestic cats in the world, with 33 different breeds. A cat's heart beats twice as fast as a human heart, at 110 to 140 beats per minuts.

25 percent of cat owners blow dry their cats hair after a bath. The largest cat breed is the Ragdoll. Males weigh twelve to twenty pounds, with females weighing ten to fifteen pounds. The smallest cat breed is the singapura. Males weigh about six pounds while females weigh about four pounds.

Overall Simulation Results

Technique	Total Chars	Recognition Rate (%)
Proposed, 3 features	3647	99.75
ABBYY	14600	96
Yasser [5], 8 features	14600	98

Conclusions and Future work

- 3 features among 8 features shows best performance in image segmentation as well as character recognition
  - Pixel (1), Mean (2) and Entropy (5)
  - Pixel (1), Mean (2) and Skewness (6)

- Two combinations shows very close results

- Future Works-
  - Image fusion using these two combinations and performance evaluation

References

1. O.D Trier and A.K. Jain, "Goal-directed evaluation of binarization methods," IEEE Trans. on Pattern Recognition And Machine Intelligence, Vol. 17, no. 12, pp. 1191-1201, 1995.
2. W. Niblack, "An introduction to Digital Image Processing", Prentice Hall, Eaglewood Cliffs, NJ, pp. 115-116, 1986.
3. P.K. Sahoo, S. Soltani,, A.K.C. Wong, "A Survey of thresholding techniques", Computer vision, Graphics and image Processing, Vol. 41, pp. 233-260, 1988
4. N. Otsu, "A Threshold Selection Method From Gray Level Histograms", IEEE Trans. On Systems, Man and Cybernetics, SMC-9, pp. 62-66, 1979
5. Y. Alginahi, "Computer Analysis of Composite Documents with Non-uniform Background", PhD Thesis, Electrical and Computer Engineering, University of Windsor, ON, Canada, 2004.
6. M.A. Sid-Ahmed, "Image Processing Theory, Algorithms and Architectures", McGraw-Hill, pp. 313-375, 1995.
7. R.C. Gonzalez, and R.E. Woods, "Digital Image Processing", Prentice-Hall, New Jersy, 2002
8. R. Koker and Y. Sari, "Neural Network Based Automatic Threshold Selection for an Industrial Vision System", Proc. Int. Conf. on Signal Processing, pp. 523-525, 2003
9. N. Papamarkos, "A Technique for Fuzzy Document Binarization", Procs. Of the ACM Symposium on 'Document Engineering, pp. 152-156, 2001.

Thanks for your Patience